Lab 3: Design & Conduct a Factorial in a RCBD

Code

AUTHOR
Angel Balbuena, Wendy Tan, Jack Vogel, Ben Cox

Introduction

Research Question and Significance

Caffeine and sugar consumption has been widely studied in relation to its effects on human health, particularly its impact on cortisol levels which are linked to stress, metabolism and inflammation regulation. By conducting this experiment, we aim to investigate how different energy drinks influence cortisol levels in a population of Islanders. Our main research question guiding this study is: How do caffeine and sugar in energy drinks affect blood cortisol levels of Islanders?

We will be using R packages from the canvas guides and some implementations from STAT 331 to conduct our setup and analysis.

Brief Literature Review

Previous research suggests that caffeine and sugar consumption can influence cortisol levels due to its concentrated sugar and caffeine content. Caffeine is known to affect cortisol levels as well as sugar, but not as much. This study aims to test this hypothesis by analyzing cortisol levels after consumption of 4 different types of energy drinks.

RCBD and Data Collection

We will design an RCBD with a total of 48 observations of islanders randomly assigned the order of the drink type as well as serving as our blocks in the experiment. To collect our data, we gave our islanders the randomly assigned drink for that day and then waited to collect their cortisol levels. We did this for all 48 but one islander (4 observations) left the study after the first day.

```
# Load in the R packages nessesary
library(tidyverse)  # For data manipulation and visualization
library(edibble)    # For experimental design and data structure handling
library(knitr)      # For meaningful displays of the data tables
library(kableExtra) # For making the tables easier to read
library(readxl)     # For reading the excel file containing our collected data
library(lme4)       # For fitting linear and generalized linear mixed-effects models
library(lmerTest)   # For adding p-values to mixed-effects models from lme4
library(emmeans)    # For comparing means of different levels of variables
library(multcomp)   # For multiple comparisons and testing hypotheses
set.seed(00000)     # seed set so data randomization is replicable
```

```
islanders <- read_csv("Islanders.csv")
set1 <- islanders |>
  dplyr::select(Name) |>
  slice(1:20) |>
  slice_sample(n = 3)
set2 <- islanders |>
  dplyr::select(Name) |>
  slice(21:40) |>
  slice_sample(n = 3)
set3 <- islanders |>
  dplyr::select(Name) |>
  slice(41:60) |>
  slice_sample(n = 3)
set4 <- islanders |>
  dplyr::select(Name) |>
  slice(61:80) |>
  slice_sample(n = 3)
islander_sample <- bind_rows(set1, set2, set3, set4)
```

RCBD Randomization

To ensure a well-structured and controlled study, we implemented a Randomized Complete Block Design (RCBD), where individual islanders served as blocks. This approach accounted for variability among participants by ensuring that each islander received all treatment combinations in a randomized order. A total of 12 islanders were randomly selected from a larger population, and each was assigned to consume four different types of energy drinks: caffeine + sugar, caffeine-only, sugar-only, and neither. The order of treatments was randomized to mitigate potential bias, with each islander receiving one treatment per day. To prevent carryover effects, a one-day washout period, equivalent to approximately 14 islander-time days or 24 hours real time, was enforced before the next treatment was administered.

```
# Create a new RCBD design and specify the structure
des <- design("Islanders RCBD") |>
  set_units(islander = c("1", "2", "3", "4", "5", "6", "7", "8", "9", "10", "11", "12"),
            drink  = nested_in(islander, 4)) |>  # Each block has 4 experimental units
  set_trts(drinktype = c("Caf/Sugar", "Caf-free/Sugar", "Caf-free/Sugar-free", "Caf/Suga
  allot_trts(drinktype ~ drink) |>  # Assign treatments to experimental units
  assign_trts("random")  # Randomly assign treatments

# Generate the experimental design table
block_table <- serve_table(des)

# Add a column for response values, initialized as NA
block_table$cortisol <- NA

# Display the design table
#kable(block_table)

# Ensure each block contains the correct number of experimental units
#block_table |>
  #count(islander) |>
  #kable()

# Ensure each treatment is replicated the correct number of times
#block_table |>
  #count(drinktype) |>
  #kable()

# Ensure each treatment x block combo is replicated only once
# (aka each treatment appears in each block once)
#block_table |>
  #count(drinktype, islander) |>
  #kable()
```

During the data collection process, each islander consumed 250 mL of the assigned energy drink, ensuring consistency in treatment administration. Since cortisol levels fluctuate throughout the day, all drinks were administered at the same time for each participant to maintain control over this variable. Following consumption, a 30-minute waiting period was observed before measuring blood cortisol levels. Blood samples were then collected and analyzed to assess cortisol responses. By carefully randomizing treatments and standardizing data collection procedures, we minimized confounding variables, allowing for a more accurate assessment of the effects of caffeine and sugar on blood cortisol levels.

Join names to RCBD

```
rcbd <- read_excel("IslandersRCBD.xlsx")
rcbd <- rcbd|>
  mutate(islander = as.character(islander)) |>
  mutate(cortisol = as.numeric(cortisol)) |>
  mutate(cortisol = round(cortisol, digits = 2))

full_table <- islander_sample |>
  mutate(islander = as.character(row_number())) |>
  full_join(rcbd, by = join_by(islander))

full_table |>
  dplyr::select(-Name.y) |>
  head(n = 8) |>
  kable(col.names = c("Name", "Islander", "Drink",
                      "Caffeine", "Sugar", "Cortisol")) |>
  kable_styling(bootstrap_options = c("striped"))
```

Name	Islander	Drink	Caffeine	Sugar	Cortisol
Dheer Mehta	1	drink01	Caf	Sugar	9.71
Dheer Mehta	1	drink02	Caf-free	Sugar	8.33
Dheer Mehta	1	drink03	Caf-free	Sugar-free	7.99
Dheer Mehta	1	drink04	Caf	Sugar-free	6.76
Natascha Eklund	2	drink05	Caf-free	Sugar	8.21
Natascha Eklund	2	drink06	Caf	Sugar-free	6.83
Natascha Eklund	2	drink07	Caf	Sugar	9.23
Natascha Eklund	2	drink08	Caf-free	Sugar-free	5.34

Statistical Methods

The first model we used was:

$$y_{ijk} = \alpha_i + \beta_j + \alpha\beta_{ij} + \rho_k + \epsilon_{ijk}$$

Where:

- i: Caffeine (1) or Caffeine-Free (2)
- j: Sugar (1) or Sugar-Free (2)
- k: The islander in question (1-11)

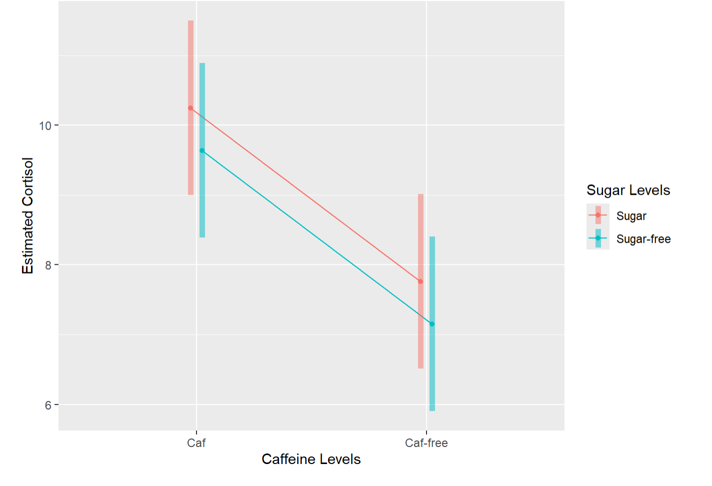
This gave us these results:

```
rcbd = rcbd |>
  filter(cortisol != "NA") # getting rid of NAs

isle_mod = lmer(cortisol ~ caffeine*sugar + (1 | islander),
                data = rcbd) # includes interaction term
anova(isle_mod)
```

```
Type III Analysis of Variance Table with Satterthwaite's method
               Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
caffeine       68.052  68.052     1    30 75.0559 1.156e-09 ***
sugar           4.081   4.081     1    30  4.5009   0.04225 *
caffeine:sugar  0.786   0.786     1    30  0.8667   0.35931
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Here we found that sugar and caffeine alone were significant, but the interaction between sugar and caffeine was not at an alpha level of 0.05. (F = 0.8667, p = 0.35931, df = (1, 30)) This lack of interaction can also be seen in the graph below. Notice how the change between caffeine levels is the same for both sugar levels.

```
# same model as before but without interaction term
isle_mod_plus = lmer(cortisol ~ caffeine + sugar + (1 | islander),
                data = rcbd)

# plot comparing differences between treatment levels of each variable
emmip(isle_mod_plus, sugar ~ caffeine, CIs = T) +
  labs(y = "Estimated Cortisol",
       x = "Caffeine Levels",
       color = "Sugar Levels")
```



Because of that we can rewrite the model without the interaction effect. That gives us the model:

$$y_{ijk} = \alpha_i + \beta_j + \rho_k + \epsilon_{ijk}$$

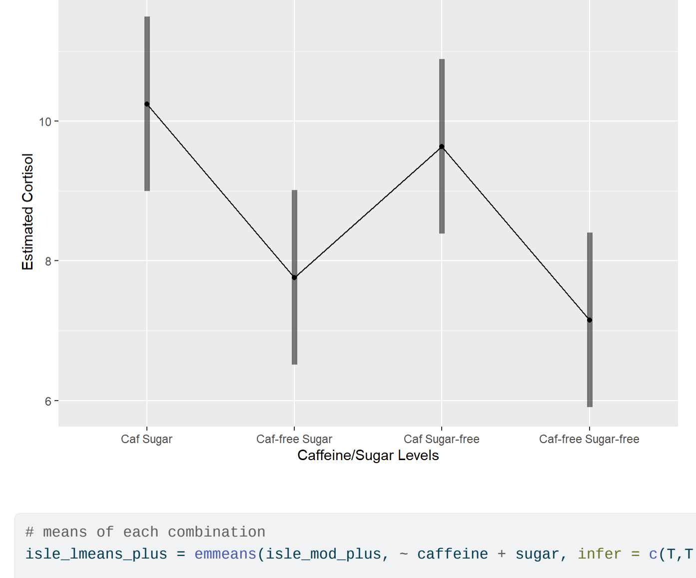
Results

```
# anova of the model that doesn't have interaction term
anova(isle_mod_plus)
```

```
Type III Analysis of Variance Table with Satterthwaite's method
         Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
caffeine 68.052  68.052     1    31 75.3802 8.357e-10 ***
sugar     4.081   4.081     1    31  4.5204   0.04156 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With the updated model and an ANOVA test can see that while both sugar and caffeine are statistically significant at an alpha level of 0.05, the effects of caffeine are far more significant. (Caffeine: F = 75.3802, p < 0.001, df = (1, 31); Sugar: F = 4.5204, p = 0.04156, df =(1, 31))

```
# plot comparing different level combos
emmip(isle_mod_plus, ~ caffeine + sugar, CIs = T) +
  labs(y = "Estimated Cortisol",
       x = "Caffeine/Sugar Levels")
```




```
# means of each combination
isle_lmeans_plus = emmeans(isle_mod_plus, ~ caffeine + sugar, infer = c(T,T))

# making it into a data table and cleaning up the data table
groups = isle_lmeans_plus |>
  cld(letters = LETTERS, decreasing = T, adjust = "sidak")
groups = as.data.frame(groups) |>
  mutate(.group = str_trim(.group),
         .group = case_when( .group == "1" ~ "A",
                             .group == "2" ~ "B",
                             TRUE ~ .group),
         p.value = if_else(as.numeric(p.value) < 0.001, "< 0.001",
                           format(as.numeric(p.value), digits = 3)),
         across(c(emmean, SE, df, lower.CL, upper.CL, t.ratio),
                ~ round(.x, digits = 3)))

# the actual table as outputted
groups |>
  arrange(desc(emmean)) |>
  kable(col.names = c("Caffeine", "Sugar", "mean", "SE", "df", "lower bound",
                      "upper bound", "t-ratio", "p-value", "group"),
        caption = "Means of Caffeine/Sugar Levels") |>
  kable_styling(bootstrap_options = c("striped")) |>
  column_spec(1:10, border_left = T, border_right = T) |>
  column_spec(6:7, width = "2cm")
```

Means of Caffeine/Sugar Levels

Caffeine	Sugar	mean	SE	df	lower bound	upper bound	t-ratio	p-value	group
Caf	Sugar	10.248	0.578	12.911	8.577	11.919	17.719	< 0.001	A
Caf	Sugar-free	9.639	0.578	12.911	7.968	11.310	16.666	< 0.001	A
Caf-free	Sugar	7.761	0.578	12.911	6.090	9.432	13.419	< 0.001	B
Caf-free	Sugar-free	7.152	0.578	12.911	5.481	8.823	12.365	< 0.001	B

```
# comparisons between the different treatment level combos
comparisons = isle_lmeans_plus |>
  pairs(adjust = 'tukey', infer = c(T,T))
comparisons = as.data.frame(comparisons) |>
  mutate(across(c(estimate, SE, df, lower.CL, upper.CL, t.ratio, p.value),
                ~ round(.x, digits = 4)))

# comparisons table output
comparisons |>
  arrange(desc(estimate)) |>
  kable(col.names = c("Contrast", "Estimated Difference", "SE", "df", "Lower Bound",
                      "Upper Bound", "t-ratio", "p-value"),
        caption = "Differences Between Caffeine/Sugar Levels") |>
  kable_styling(bootstrap_options = c("striped")) |>
  column_spec(1:7, border_left = T, border_right = T) |>
  column_spec(2, width = "2em")
```

Differences Between Caffeine/Sugar Levels

Contrast	Estimated Difference	SE	df	Lower Bound	Upper Bound	t-ratio	p-value
Caf Sugar - (Caf-free Sugar-free)	3.0964	0.4051	31	1.9968	4.1960	7.6426	0.0000
Caf Sugar - (Caf-free Sugar)	2.4873	0.2865	31	1.7097	3.2648	8.6822	0.0000
(Caf Sugar-free) - (Caf-free Sugar-free)	2.4873	0.2865	31	1.7097	3.2648	8.6822	0.0000
Caf Sugar - (Caf Sugar-free)	0.6091	0.2865	31	-0.1684	1.3866	2.1261	0.1673
(Caf-free Sugar) - (Caf-free Sugar-free)	0.6091	0.2865	31	-0.1684	1.3866	2.1261	0.1673
(Caf-free Sugar) - (Caf Sugar-free)	-1.8782	0.4051	31	-2.9778	-0.7786	-4.6358	0.0003

Conclusion

This study examined the effects of caffeine and sugar on blood cortisol levels through a randomized factorial experiment. Our results demonstrated that both caffeine and sugar significantly impacted cortisol levels, with caffeine having a much stronger effect (F = 75.3802, p < 0.001) than sugar (F = 4.5204, p = 0.04156). While the interaction between caffeine and sugar was not statistically significant (F = 0.8667, p = 0.35931), our findings suggest that caffeine alone plays a dominant role in increasing cortisol levels. Given the physiological and psychological effects of high cortisol—such as weight gain, fatigue, and memory issues—these results underscore the importance of understanding how dietary choices, particularly caffeine intake, influence stress hormone regulation. Further studies could explore additional factors influencing cortisol levels, such as habitual caffeine consumption, stress levels, or sleep quality, to refine our understanding of the broader implications.